A common mass scale for satellite galaxies of the Milky Way

Louis E. Strigari[1], James S. Bullock[1], Manoj Kaplinghat[1], Joshua D. Simon[2], Marla Geha[3], Beth Willman[4], Matthew G. Walker[5]

[1]*Center for Cosmology, Department of Physics and Astronomy, University of California, Irvine, CA 92697-4574, USA*
[2]*Department of Astronomy, California Institute of Technology, 1200 E. California Blvd., MS105-24, Pasadena, CA 91125, USA*
[3]*Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520-8101, USA*
[4]*Harvard-Smithsonian Center for Astrophysics, 60 Garden St. Cambridge, MA 02138, USA*
[5]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*

The Milky Way has at least twenty-three known satellite galaxies that shine with luminosities ranging from about a thousand to a billion times that of the Sun. Half of these galaxies were discovered [1,2] in the past few years in the Sloan Digital Sky Survey, and they are among the least luminous galaxies in the known Universe. A determination of the mass of these galaxies provides a test of galaxy formation at the smallest scales [3,4] and probes the nature of the dark matter that dominates the mass density of the Universe [5]. Here we use new measurements of the velocities of the stars in these galaxies [6,7] to show that they are consistent with them having a common mass of about 10^7 M$_\odot$ within their central 300 parsecs. This result demonstrates that the faintest of the Milky Way satellites are the most dark matter-dominated galaxies known, and could be a hint of a new scale in galaxy formation or a characteristic scale for the clustering of dark matter.

Many independent lines of evidence strongly argue for the presence of dark matter in galaxies, clusters of galaxies, and throughout the observable Universe [5]. Its identity, however, remains a mystery. The gravity of dark matter overwhelms that of the normal atoms and molecules and hence governs the formation and evolution of galaxies and large-scale structure [8–10]. In the currently favored dark matter models, structure in the Universe forms hierarchically with smaller gravitationally bound clumps of dark matter – haloes– merging to form progressively larger objects.

The mass of the smallest dark matter halo is determined by the particle properties of dark matter. Dark matter candidates characterized as cold dark matter can form haloes that are many orders of magnitude smaller than the least luminous haloes that we infer from observations. Cosmological simulations of cold dark matter predict that galaxies like the Milky Way should be teeming with thousands of dark matter haloes with masses $\sim 10^6 M_\odot$, with a steadily increasing number as we go to the smallest masses [11–14]. A large class of dark matter candidates characterized as "warm" would predict fewer of these small haloes [15]. However, even for cold dark matter it is uncertain what fraction of the small dark matter haloes should host visible galaxies, as the ability

of gas to cool and form stars in small dark matter haloes depends on a variety of poorly-understood physical processes [16–20].

The smallest known galaxies hosted by their own dark matter haloes are the dwarf spheroidal satellites of the Milky Way [3,4]. These objects have very little gas and no signs of recent star formation. The least luminous galaxies were recently discovered in the Sloan Digital Sky Survey (SDSS) [1,2] and follow-up observations have revealed them to be strongly dark matter dominated [6,21,22].

We have compiled line-of-sight velocity measurements of individual stars in 18 of the 23 known dwarf galaxies in the Milky Way [6,7]. We use these measurements to determine the dynamical mass of their dark matter haloes using a maximum likelihood analysis [23]. The dynamical mass is best constrained within the stellar extent, which corresponds to an average radius of approximately 0.3 kiloparsecs (kpc) for all the satellites. We determine this mass, $M_{0.3}$, by marginalizing over a five-parameter density profile for dark matter that allows for both steep density cusps and flat cores in the central regions. It is important to note that the observed velocity dispersion of stars is determined by both the dynamical mass and the average anisotropy of the velocity dispersion (that is, difference between tangential and radial dispersion). The anisotropy is unknown and hence we marginalize over a three-parameter stellar velocity anisotropy function that allows us to explore a range of orbital models for the stars [23].

Figure 1 shows the resulting determination of $M_{0.3}$. We find that all 18 dwarf galaxies are consistent with having a dynamical mass of 10^7 solar masses within 0.3 kpc of their centre, despite the fact that they have luminosity differences over four orders of magnitude. This result implies a dark matter central density of $\sim 0.1\,M_\odot/pc^3$ in these galaxies. Earlier studies suggested that the highest-luminosity dwarf galaxies all shared a common mass [4,24]. With larger stellar data sets, more than double the number of dwarf galaxies, and more detailed mass modeling, our results confirm this earlier suggestion and conclusively establish that the dwarf galaxies of the Milky Way share a common mass scale.

Because of the proximity of the dwarf galaxies to the Milky Way, it is possible that tidal effects could change the velocities of stars and thus affect the mass measurements. In the kinematic data, tidal forces could be revealed as a velocity gradient across the observed plane of the dwarf [25,26]. We have tested the dwarf galaxies for velocity gradients and have found no conclusive evidence of tidal effects (see Supplementary Information).

We fitted a $M_{0.3}$-luminosity relation to the data and obtained $M_{0.3} \propto L^{0.03\pm0.03}$. This result does not change significantly if we use the luminosity contained within 0.3 kpc rather than the total luminosity. The common mass scale of $\sim 10^7 M_\odot$ may thus reflect either a plummeting efficiency for galaxy formation at this mass scale, or that dark matter haloes with lower masses simply do not exist.

The characteristic density of $0.1\,\mathrm{M_\odot/pc^3}$ may be associated with a characteristic halo formation time. In theories of hierarchical structure formation, the central density of dark matter haloes is proportional to the mean density of matter in the Universe when the halo formed. The earlier the formation, the higher the density. For cold dark matter models, our measurement implies that these haloes collapsed at a redshift greater than about 12, or earlier than 100 million years after the Big Bang. Measurements of the cosmic microwave background [27] suggest that the Universe went from being neutral to ionized at redshift 11 ± 1.4. These dark matter haloes thus formed at approximately the same time that the Universe was re-ionized.

Within the context of the cold dark matter theory, high-resolution cosmological simulations can be used to relate the mass within the central regions of the dark matter halo to the depth of the gravitational potential well [28]. Simulations show that $\mathrm{M_{0.3}} \approx 10^7 \mathrm{M_\odot}\,(\mathrm{M_{total}}/10^9 \mathrm{M_\odot})^{0.35}$, where $\mathrm{M_{total}}$ is the mass of the halo before it was accreted into the Milky Way host potential. Thus, it is possible that the implied total mass scale of $10^9 \mathrm{M_\odot}$ reflects the characteristic scale at which supernova feedback [29] or the imprint of the re-ionization of the universe [18,30] could sharply suppress star formation.

Perhaps a more speculative, but certainly no less compelling, explanation of the common mass scale is that dark matter haloes do not exist with $\mathrm{M_{0.3}}$ below $\sim 10^7\,\mathrm{M_\odot}$. This implies that these dwarf galaxies inhabit the smallest dark matter haloes in the Universe. Warm dark matter particles have larger free streaming length than standard cold dark matter particles, which implies that density perturbations are erased below a characteristic length scale, resulting in a higher minimum mass for dark matter haloes. A thermal warm dark matter candidate with mass of approximately $1\,\mathrm{keV}$ would imply a minimum halo mass of $10^9 \mathrm{M_\odot}$. Thus, our mass determinations rule out thermal warm dark matter candidates with masses less than about $1\,\mathrm{keV}$, but dark matter masses somewhat larger than 1 keV would yield a minimum dark matter halo mass consistent with the mass scale we observe.

Future imaging surveys of stars in the Milky Way will provide a more complete census of low-luminosity Milky Way satellites, with the prospects of determining whether astrophysics or fundamental dark matter physics is responsible for setting the common mass scale. In particular, the masses for the faintest dwarf galaxies will become more strongly constrained with more line-of-sight velocity data. This will sharpen the observational picture of galaxy formation on these small-scales and provide data around which theories of galaxy formation may be built.
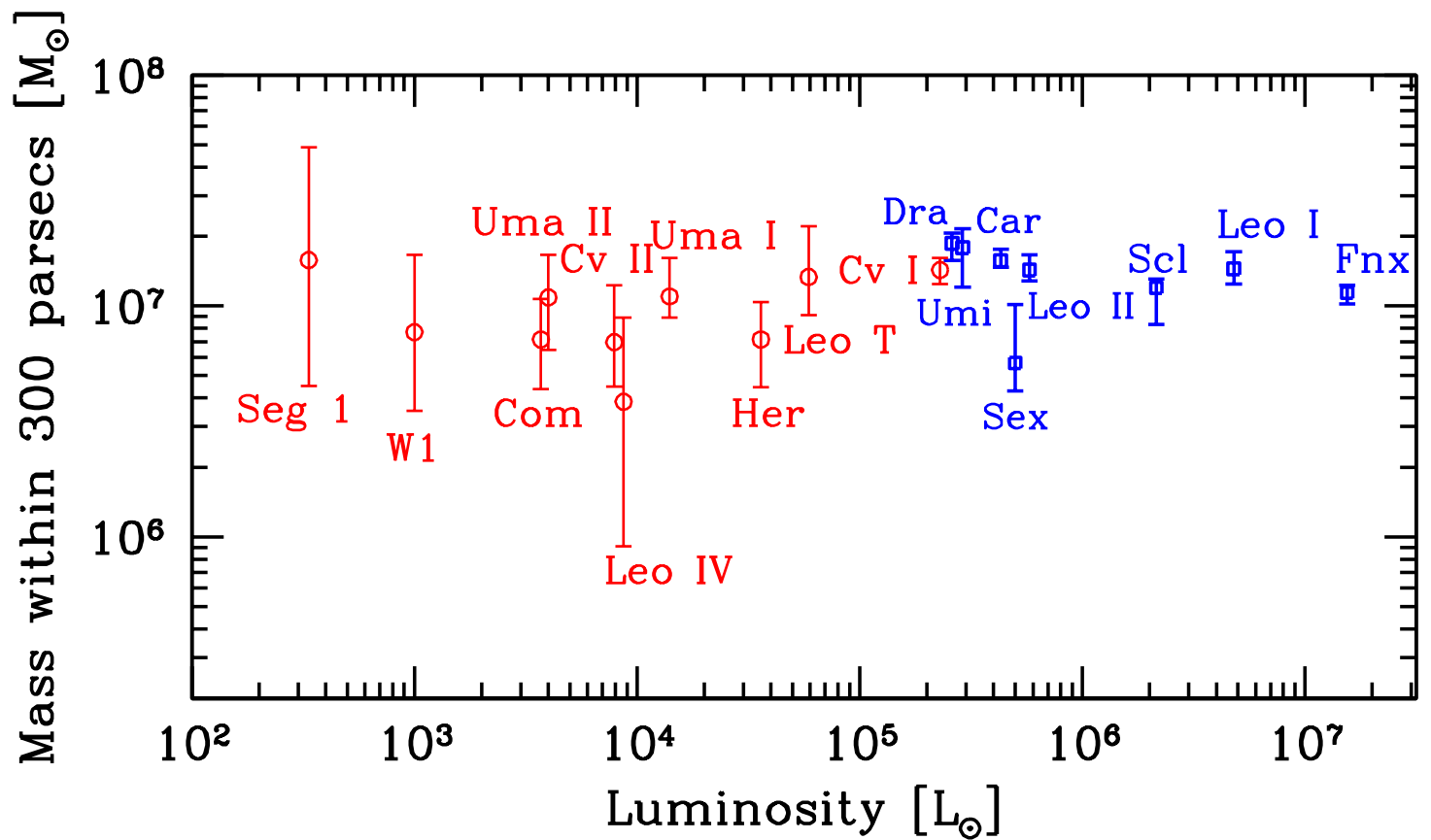
Figure 1: The integrated mass of the Milky Way dwarf satellites, in units of solar masses, within their inner 0.3 kpc as a function of their total luminosity, in units of solar luminosities. The circle (red) points on the left refer to the newly-discovered SDSS satellites, while the square (blue) points refer to the classical dwarf satellites discovered pre-SDSS. The error bars reflect the points where the likelihood function falls off to 60.6% of its peak value.

4

Supplementary Information

In this section we present the details of our dynamical mass modeling. We begin by reviewing the Jeans formalism for determining the mass of a galaxy using line-of-sight velocities. We then discuss our specific treatment of the line-of-sight velocity data and introduce the likelihood function that we use in our analysis. Next, we present our results and discuss the implications. Finally, we discuss the systematics that may affect the mass modeling.

1 Mass modeling

When modeling the stellar distribution of a dwarf galaxy in close proximity to the Milky Way (MW), the internal gravitational force from the dwarf must be compared to the external force from the Milky Way. Generally, the internal gravitational force from the dwarf is $\sim \sigma^2/R_s$, where σ is the velocity dispersion and R_s is the half-light radius of the dwarf. The external tidal force from the MW potential is $\sim (220\,\mathrm{km/s})^2 R_s/D^2$, where D is the distance to the dwarf from the center of the MW, and $220\,\mathrm{km/s}$ is the approximate rotation speed of the MW in the outer regions of the halo where the dwarfs are located. The most luminous of the MW dwarfs have half-light radii of $R_s \sim 400$ pc, and the least luminous of the dwarfs have half-light radii of $R_s \sim 10 - 100$ pc. The velocity dispersions vary in the range $\sigma \sim 5 - 15$ km s^{-1}. Comparing the internal and external forces on dwarfs in the observed distance range of $\sim 20 - 250$ kpc, we find that the internal gravitational forces are typically larger by ~ 100. Note that this estimate does not exclude the possibility that the dwarfs have been tidally stripped in the past; it does, however, allow us to proceed safely with the assumption that the surviving stellar distributions trace the local potential.

Previous studies of some Milky Way dwarf spheroidals (dSphs) show that these objects exhibit no streaming motions [31,32]. Making the further assumption that these systems are in steady state, the three-dimensional radial velocity dispersion of the system, σ_r, is given by the Jeans equation

$$r\frac{d(\rho_\star \sigma_r^2)}{dr} = -\rho_\star(r)\frac{GM(r)}{r} - 2\beta(r)\rho_\star \sigma_r^2. \tag{1}$$

This equation is valid if the potential of the system is spherically-symmetric, which has been shown to be a good description of dark matter satellites in a host potential [33]. To convert to the observable quantity, we must integrate the solution to the Jeans equation along the line-of-sight. Performing this integration gives [34]

$$\sigma_t^2(R) = \frac{2}{I_\star(R)} \int_R^\infty \left(1 - \beta\frac{R^2}{r^2}\right) \frac{\rho_\star \sigma_r^2 r}{\sqrt{r^2 - R^2}} dr. \tag{2}$$

Here R is defined as the projected radius in the plane of the sky. The three-dimensional stellar density profile is defined by $\rho_\star(r)$, which is determined from the projected stellar distribution, $I_\star(R)$. The stellar velocity anisotropy is defined as $\beta = 1 - \sigma_\theta^2(r)/\sigma_r^2(r)$, where σ_θ is the tangential

component of the velocity dispersion. We now discuss in turn our parameterizations of the different functions entering in equation 2.

Stellar Surface Density It is standard to fit the stellar surface densities of the systems we study to either Plummer or King profiles. The surface density for the King profile is [35]

$$I_{\text{king}}(R) = k \left[\left(1 + \frac{R^2}{r_c^2} \right)^{-1/2} - \left(1 + \frac{r_{lim}^2}{r_c^2} \right)^{-1/2} \right]^2, \tag{3}$$

which results in a de-projected three-dimensional density of

$$\rho_{\text{king}}(r) = \frac{k}{\pi r_c [1 + (r_{lim}/r_c)^2]^{3/2} z^2} \left[\frac{1}{z} \cos^{-1} z - \sqrt{1 - z^2} \right], \tag{4}$$

where $z^2 = (1 + r^2/r_c^2)/(1 + r_{lim}^2/r_c^2)$. The normalization constant, k, for the King profile thus is irrelevant when applying the Jeans equations. The King profile depends on two parameters, r_{lim} and r_c.

The surface density for the Plummer profile is given by

$$I_{\text{pl}}(R) = \frac{4}{3} \frac{\rho_0 r_{\text{pl}}}{[1 + (R/r_{\text{pl}})^2]^2}, \tag{5}$$

which results in a de-projected three-dimensional density of

$$\rho_{\text{pl}}(r) = \frac{\rho_0}{[1 + (r/r_{\text{pl}})^2]^{5/2}}. \tag{6}$$

The only relevant free parameter in the Plummer profile is r_{pl}.

In Table 1 we show the respective fits to the surface density for each of the dwarf satellites [6,36]. Many of the well-known dSphs are well-fit by King profiles, although for some galaxies the King profile fits have been updated to account for the observed distribution of stars in the outer regions [21]. As is seen, the majority of the new satellites are well-fit by Plummer profiles. In the instances where both Plummer and King profiles have been fit to the data, we find that the exact form of the fit does not strongly affect the results we present below.

Velocity Anisotropy We assume that both tangential components of the velocity dispersion are equal, $\sigma_\theta^2 = \sigma_\phi^2$. Many of the systems we study are observed to have multiple stellar populations [37], so there is no reason why the anisotropy should be constant throughout the galaxy. To account for radial variation, we parametrize the anisotropy profile as

$$\beta(r) = (\beta_\infty - \beta_0) \frac{r^2}{r_\beta^2 + r^2} + \beta_0. \tag{7}$$

The velocity anisotropy profile of this form is thus described by an asymptotic inner value, β_0, and asymptotic value near the edge of the halo, β_∞, and a scale radius, r_β. We place the constraints on β_0 and β_∞ such that $\beta(r) < 1$ for all radii.

Dark Matter Density Profile To model the dark matter mass distribution, we use a density profile of the form

$$\rho(r) = \frac{\rho_0}{(r/r_0)^a[1 + (r/r_0)^b]^{(c-a)/b}}. \tag{8}$$

The asymptotic inner slope is determined by a, the asymptotic outer slope is determined by c, and the transition between these two regimes is determined by b. The scale density is defined as ρ_0, and the scale radius is defined as r_0. In all cases, the line-of-sight velocities are not able to determine the parameters a, b, c, though as we see below, the data do strongly constrain the mass and the density of these systems at a characteristic radius [23].

For some regions of parameter space, the density profile in equation 8 has an infinite mass, so it cannot represent a physically reasonable dark matter halo. To ensure that the mass is finite in all of parameter space, we weight equation 8 as $\rho(r) \rightarrow \rho(r) \exp^{-r/r_{\text{cut}}}$. We determine the cut-off radius, r_{cut}, from the standard Roche-limit criteria,

$$r_{\text{cut}} \simeq \left(\frac{GM_{\text{halo}}D^2}{2\sigma_{MW}^2} \right)^{1/3}, \tag{9}$$

where M_{halo} is the total mass of the satellite, σ_{MW} is the velocity dispersion of the Milky Way at the position of the satellite, and D is the distance to the center of the satellite. To establish a conservative upper limit on r_{cut}, in particular allowing it to span the largest physically possible range, for all of the satellites we take $M_{\text{halo}} = 10^9 \, M_\odot$ and $\sigma_{\text{MW}} = 200 \text{ km s}^{-1}$.

2   Data Modeling

From a data set consisting of line-of-sight velocities, our goal is to determine the constraints on any of the parameters in equation 2. To evaluate the constraints on these parameters, we undertake a maximum likelihood analysis and construct the probability distribution for obtaining a set of line-of -sight velocities in a dwarf galaxy. The goal of this section is to present and discuss our implementation of the maximum likelihood analysis.

Likelihood Function The dispersion in the velocity at a given position is the sum of two components: 1) the dispersion from the velocity distribution function, and 2) the random error stemming from the uncertainty in the measurement of the velocity of each star. Motivated by theoretical modeling [38] and observations of line-of-sight velocities [6], we take both of these distributions to be Gaussian. The dispersion stemming from the velocity distribution function is determined by the set of parameters in equations 7 and 8 that govern the dark matter halo model and the velocity anisotropy. We refer to this vector set of parameters as $\vec{\theta}$. The probability for obtaining a set of line-of-sight velocities, $\mathbf{v}$, given the theoretical parameters, is $P(\mathbf{v}|u, \sigma_{\mathbf{t}})$. Here u is defined as the systemic velocity of the galaxy, and $\sigma_{\mathbf{t}}$ is determined from equation 2. It is thus implied that $\sigma_{\mathbf{t}}$ is a function of the parameters $\vec{\theta}$. With the above assumptions, the result for $P(\mathbf{v}|u, \sigma_{\mathbf{t}})$ is a Gaussian

distribution,

$$P(\mathbf{v}|u, \sigma_{\mathbf{t}}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2}\frac{(v_i - u)^2}{\sigma_i^2}\right]. \tag{10}$$

The product is over the N number of stars in the galaxy with line-of-sight velocity measurements. The total variance at the projected radius R (at the position of the ith star) is thus given by $\sigma_i^2 = \sigma_{t,i}^2 + \sigma_{m,i}^2$, where $\sigma_{t,i}^2$ refers to the variance from the theoretical distribution, and $\sigma_{m,i}^2$ is the variance from the measurement uncertainty. In writing equation 10 we have assumed no correlations between both the theory and measured dispersions. Since we are assuming no streaming motion, $\sigma_{t,i}$ is replaced with equation 2. According to Bayes theorem, equation 10 is proportional to the probability of the parameters given the data, i.e. $P(\mathbf{v}|u, \sigma_{\mathbf{t}}) \propto P(u, \sigma_{\mathbf{t}}|\mathbf{v})$. When considered as a function of the parameters, equation 10 can be then defined as the likelihood function for the parameters, $\mathcal{L}(\vec{\theta})$.

It is important to note that the assumption of Gaussianity for the intrinsic velocity distribution function is only an approximation to the true phase space distribution. The true intrinsic distribution function will depend on the exact form of the density of the stellar population and the potential of the background dark matter halo. Taking the stellar distribution to be a King profile, the Gaussian approximation provides a good estimation of the true distribution, though there may be some deviations from Gaussianity in the outer most regions of galaxies [31]. From an observational perspective, the distribution in velocities for nearly all of the dwarf satellites we consider is well-described by a Gaussian distribution at the $\sim 2 - 3\sigma$ level. For all satellites, we have found that tests of deviation from Gaussianity in the data have proved inconclusive.

In addition to deviations due to the nature of the local gravitational potential, from a systematic perspective the observed velocity distribution may deviate from Gaussianity if there is large contamination to the sample from interloping stars or from stars not bound to the galaxy because of tidal interactions. To accurately model the mass distribution of a galaxy with equations 1 and 2 and line-of-sight velocities, interloping and unbound stars must be properly identified. Positional and photometric criteria provide a means for identifying these interloping and unbound stars [6]. Numerical simulations have also been performed to aid in the identification of contaminating stars [25,38]. Tagging stars as outliers based simply on their presence in the tails of the velocity distribution introduces an intrinsic bias into the determination of the mass, so excluding these stars as non-members must be done with caution. Once interlopers are cut, it is observed that the mass determined from the full-data sets is accurate to $\sim 20\%$ [25,38]. This error is most relevant, however, in determining the *total* halo mass. The presence of interlopers is even less significant when determining the mass at radii nearer to the center of the halo, as we do below.

3 Results

We now turn to application of the analysis above and determine the best fitting masses for the Milky Way dwarf satellites. The primary result will be integrated masses within radii of 0.1 and 0.3 kpc. The latter provide the main results of our analysis. The former is convenient because, for a few satellites, it requires less of an extrapolation of the dark matter halo beyond the observed stellar distribution. Our choice of 0.3 kpc as the characteristic radius probes the mass further in the interior of the haloes than the characteristic radius of 0.6 kpc used in a previous study [23]. This latter choice of 0.6 kpc was better suited for comparing the masses of only the pre-SDSS, classical dSphs, which are on average more extended than the new SDSS population. For dwarf satellites with over a hundred line-of-sight velocities, we find the strength of the constraints on the mass within 0.3 and 0.6 kpc to be typically similar.

The mass within a given radius, m, is obtained by integrating the probability distribution in equation 10 over the model parameters,

$$\mathcal{L}(m) \propto \int P[\mathbf{v}|u, \sigma_{\mathbf{t}}(\vec{\theta})]\delta(m - M)d\vec{\theta}. \tag{11}$$

In practice, determining $\mathcal{L}(m)$ then requires an integration over all of the free parameters in equations 7 and 8. In equation 11, M is determined by the halo model parameters at the given point in parameter space. Note also that in equation 11 we have ignored the normalization of the likelihood, which does not depend on the model parameters. To ease in quoting confidence limits, we simply normalize the peak of the resulting integrated likelihood function to unity.

We assume uniform priors on the model parameters over their respective ranges. We have tested the assumption of uniform priors by also considering Gaussian priors on each parameter over their respective allowed ranges. For either assumed prior, our mass constraints are found to be robust. For the dark matter halo parameters, uniform priors are chosen over the following ranges: $0.01 < r_0 < 10$ kpc, $0 < a < 1.5$, $0.5 < b < 1.5$, and $2 < c < 4$. We find that all of these parameters are not well-constrained by the line-of-sight velocity data. We also include a cut-off radius in the halo profile as in equation 9, which, as discussed above, serves to makes the truncation of the halo sharper than the power law behavior implied by the outer slope c.

For the radially-variable anisotropy model, the following ranges are taken: $0.1 < r_\beta < 10$, $-10 < \beta_\infty < 1$, and $-10 < \beta_0 < 1$. The choice of a flat prior on the velocity anisotropy parameters does in principle give a stronger weight to more tangential orbits, though we find that including a prior on the anisotropy that more uniformly weighs radial orbits does not affect the results that we present. Though we have assumed a radially-variable anisotropy profile in equation 7, we find that our resulting constraints on the mass are equivalent to those obtained in a model where the velocity anisotropy is constant in radius. The only instance in which our mass constraints weaken significantly is in the physically unlikely scenario that the anisotropy profile has a very steep transition right around the King (or Plummer) radius.

9

Figure 2 depicts the main results of our analysis: the mass likelihood functions within 0.3 kpc for 18 of the Milky Way dwarfs. The classical, pre-SDSS dwarfs, which have more measured line-of-sight velocities spread over a larger radial distance, have likelihoods that are much more strongly constrained. Figure 2 shows that the $M_{0.3}$ values are tightly bunched around $\sim 10^7$ $M_\odot$. The only dwarf that displays a large tail in its likelihood at $\sim$ 1-σ to small $M_{0.3}$ is Leo IV, which may be related to the shorter integration time spent on Leo IV as well as the relatively small sample of stellar velocities in this galaxy.

The stellar luminosities for the Milky Way dwarf satellites span a range of approximately four orders of magnitude: from the most luminous satellite (Fornax, $\sim 10^7$ $L_\odot$), to the least luminous (Willman 1, Ursa Major II, Coma Berenices, and Segue 1: $\sim 10^3$ $L_\odot$). Since typical stellar mass-to-light ratios are of order unity, we can immediately deduce that all of these dwarfs are strongly dark matter dominated within the limiting radius of their stellar distributions. Nearly all of these dwarfs have mass-to-light ratios within their inner 0.3 kpc of $M_{0.3}/L_\odot > 10$. The most luminous systems, including Fornax, Sculptor, and Leo I, have the smallest total $M_{0.3}/L_\odot$. Because our analysis determines the total dynamical mass, the stars in these galaxies likely make a significant contribution to the measured $M_{0.3}$. However, when casting the mass-to-light ratio in terms of total halo mass, M_{total}, even for these most luminous dwarfs we have $M_{\mathrm{total}}/L > 100$ [23].

In Table 1, we show the numerical results for the mass within 0.1 and 0.3 kpc for each dwarf satellite, along with other various observed properties. The results of Table 1 also show that the masses within 100 pc are well-constrained, implying a common mass of $\sim 10^6$ $M_\odot$ within this radius. As mentioned above, for several dwarf satellites, in particular Segue 1, Willman 1, and Coma Berenices, determining the mass within 0.1 kpc requires less of an extrapolation beyond the limiting radius of the stellar distributions. In these instances, the mass within 0.1 kpc is in good agreement with masses determined from a single-component, mass-follows-light analysis [6, 21, 22, 39].

Of the newly-discovered satellites, we have not displayed results for either Bootes I or Bootes II. Line-of-sight velocities for Bootes I have been published [21,22], though for ease of comparison we have chosen to use common data sets. Our initial estimate does in fact also show that Bootes I is consistent with the $M_{0.3} \simeq 10^7$ $M_\odot$ mass scale at $\sim 2 - \sigma$. We are currently unable to analyze Bootes II because it was discovered within the past year [40] and does not yet have measured line-of-sight velocities. Of the classical, pre-SDSS satellites, we do not determine the mass for the Large Magellanic Cloud (LMC), Small Magellanic Cloud (SMC), or Sagittarius. It is likely that both the LMC and SMC have significant baryonic contribution to the their respective masses within 0.3 kpc, so a detailed determination of the halo mass of each galaxy will require accurate modeling of this contribution. Sagittarius has published line-of-sight velocities [41], though because it is in the process of tidal disruption, streaming motions will strongly affect the determination of its mass, making the above analysis not reliable.

As discussed above, for the respective stellar distributions, we use the King or Plummer

profiles given in Table 1. For several of the classical dwarfs, it has been found that in the outer regions of the galaxy a faint population of stars exists that has a surface density that falls off less steeply than either the King or Plummer profile [21, 42]. Interpreting these stars as bound to the galaxy would require an increase in the total halo mass. However, we find that the presence of these stars has little effect on the mass within 0.3 kpc, as this latter quantity is primarily determined by the population of stars that are well-described by the respective parameters in Table 1.

It is remarkable to note that, when considering both the wide range of scale radii in Table 1 and the range of velocity dispersions, the common mass scale persists. Though determination of the masses requires a numerical solution to the Jeans equation as described above, we can gain some insight into the scaling of the mass with both the velocity dispersion and the scale radius by considering the limit in which $\beta = 0$, the surface density of stars is given by a Plummer profile, and the velocity dispersion is constant as a function of radius. The latter simplification is a good approximation for many of the dwarfs. Under these assumptions, the mass at any radius is given by $M(r) \simeq 1.2 \times 10^3 \ \mathrm{M_\odot} \left[\frac{r}{\mathrm{pc}} \right] \left[\frac{\sigma}{\mathrm{km/s}} \right]^2 \frac{y^2}{1+y^2}$, where $y = r/r_{pl}$. Examining this formula, we see that, for a fixed projected velocity dispersion, the mass will *increase* if the Plummer scale radius is reduced. Qualitatively, this scaling can be understood by noting that it takes a deeper potential well to confine stars with a larger dispersion to a small radius. We note that all of our results are consistent with the minimum mass obtained from the virial theorem [43]; for example for Willman 1 we obtain a minimum mass of $\sim 3 \times 10^5 \ \mathrm{M_\odot}$, and for Fornax we obtain a minimum mass of $\sim 2 \times 10^6 \ \mathrm{M_\odot}$.

It should be stressed again that in our dynamical mass analysis we have assumed that the stars are orbiting in a spherically symmetric potential. This is a reasonable assumption because numerical simulations show [33] that potentials of these dark matter haloes are close to spherical with triaxial axis ratios of 0.8 to 0.9. Dispersion and bias introduced in the dynamical mass determinations at the level of 10-20% are not important at the present stage but will be as data sets get larger for the fainter satellites.

4 Tidal Disruption, Rotation, and Binary Contamination

Equation 10 assumes that rotational motion or external tidal forces do not contribute to the observed line-of-sight velocities. However, this may not be a complete description of the dwarf satellites, as it is very likely that their dark matter haloes have been affected by tidal interactions in the past. Although the simple scalings above indicate that it is unlikely tidal forces dominate the dynamics of these systems, it is important to quantify any deviations from steady-state dynamics more precisely. The goal of this section is to use the line-of-sight velocity data to provide conservative tests for both rotation and tidal disruption.

Tidal Disruption Tidal forces induce a gradient in the line-of-sight velocities across a galaxy [25, 26]. To test for a velocity gradient, we describe the stars in terms of their projected radial distance, R,

from the center of the dwarf, and the azimuthal angle of each star, ϕ, relative to a fixed coordinate system. Streaming motion would thus appear as a velocity gradient about an angle ϕ_0, which is determined from the data [44]. In the presence of streaming, the systemic velocity in the likelihood function of equation 10 is replaced as

$$u \rightarrow u + A \sin(\phi_i + \phi_0). \tag{12}$$

Here A represents the amplitude of the streaming motion, ϕ_0 is the (projected) axis of the streaming motion, and ϕ_i is the azimuthal angle of the ith star. It is important to note that, in principle, replacing the systemic velocity with equation 12 provides only an approximation to streaming motion that serves to pick out a dipolar term in the velocity field. Higher order multipoles may exist as a result of tidal forces; however for our purposes we assume these higher order terms to be sub-dominant to the leading dipolar term given in equation 12.

A non-zero detection of A suggests the presence of tidal forces or rotational motion. We have examined the line-of-sight velocities for each system and have determined the likelihood for the amplitude A after marginalizing over all of the halo parameters, anisotropy parameters, as well as ϕ_0. In Figure 3, we show the resulting likelihood functions for A for three systems: Willman 1, Coma Berenics, and Ursa Major II. We show these systems as examples because they are the amongst the nearest to the Milky Way, and thus may be plausible candidates for tidal disruption. As is seen, there is no significant detection of A. For the Willman 1 sample we use, the mean value of A is ~ 2 km s^{-1}, but this object is still consistent with no rotation at $\sim 1 - \sigma$.

In all of the remaining newly-discovered satellites, we find no statistically significant detection of A. In nearly all cases, we place strong upper limits on A to be a fraction of the intrinsic dispersion of each system. We note that the inconclusive detection of velocity gradients in Coma Berenices and Ursa Major II differ from the results reported in Simon and Geha [6], where a weighted mean of the velocities on each side of the position angle showed an apparent velocity gradient. Here we find that accounting for correlations with higher order velocity moments washes out the detection of a velocity gradient in each of these systems. These results imply that rotational motion or tidal disruption in the above parametrization likely does not significantly affect the mass modeling.

We note that, when interpreting the parametrization in equation 12 as a rotational signal, it is only an approximation to the true three-dimensional rotation. To construct a more physically-plausible model one would need to replace equation 12 with a model that depends on at least two angles that describe both the inclination of the rotation axis and the line-of-nodes of the system. Additionally, flattening of the stellar distribution must be accounted for by modifying the spherically-symmetric Jeans equation. At the very least, these effects will increase the uncertainty on the rotation amplitude, so in this sense our errors on the rotation amplitude are likely too strict. Our initial estimates show that, when adding an additional angle to account for three-dimensional rotation, the error on the amplitude of the rotational motion increases by a factor of about two.

Perspective Rotation As the satellites are extended objects on the sky, the line-of-sight velocity will vary as we move across the object. As we move farther from the center of the system, the tangential motion of the object will contribute to the measured line-of-sight velocity. This effect is known as "perspective rotation," and the net result is that a non-rotating object will appear to have a velocity gradient across the system [45].

Perspective rotation can be simply parametrized by replacing the systemic velocity in the likelihood function as $u \rightarrow u + v_x x/D + v_y y/D$, where x and y are the projected positions on the sky and v_x and v_y are the tangential motions in these directions. For the purposes of our present analysis, we are interested in determining if the addition of the perspective rotation term will alter our mass estimates derived above. We find that, in all data sets, the masses we determine are robust even after accounting for perspective rotation. A full analysis of a given data set will require addition of both the streaming motion term in equation 12 and the perspective rotation term; adding both of these will result in further degeneracies that will make each separate effect more difficult to extract.

Contamination from Binary Stars A final effect we cannot include at present that may introduce a systematic in determining the masses is the contribution of internal binary star motion to the velocity dispersion. For binaries with orbits of order 100 AU, the relative binary speed is comparable to the intrinsic velocity dispersion of the satellite galaxy. Thus the bias due to binaries depends on the fraction of stars that are binaries with orbits smaller than about 100 AU for the stellar population whose velocities are being measured. For two of the most luminous satellites, Draco and Ursa Minor (which also have a higher velocity dispersion), a significant effect of internal binary motion on the measured velocity dispersion has been shown to be unlikely [46]. A detailed study for faint dwarfs will require repeat observations of stars a year or two apart; this should be available in the near future. Here we simply note that if a significant fraction of the measured dispersion was due to binary motion, then the distribution of velocities would deviate significantly from the observed Gaussian distribution.

Satellite	Distance [kpc]	r_{king}, r_{pl} [kpc]	r_{lim} [kpc]	Luminosity [10^6 L$_\odot$]	M$_{0.1}$ [10^7 M$_\odot$]	M$_{0.3}$ [10^7 M$_\odot$]
Ursa Minor (Umi)	66	0.30	1.50	0.29	$0.21^{+0.09}_{-0.14}$	$1.79^{+0.37}_{-0.59}$
Draco (Dra)	80	0.18	0.93	0.26	$0.09^{+0.20}_{-0.02}$	$1.87^{+0.20}_{-0.29}$
Sculptor (Scl)	80	0.28	1.63	2.15	$0.15^{+0.28}_{-0.10}$	$1.20^{+0.11}_{-0.37}$
Sextans (Sex)	86	0.40	4.00	0.50	$0.06^{+0.02}_{-0.01}$	$0.57^{+0.45}_{-0.14}$
Carina (Car)	101	0.26	0.85	0.43	$0.48^{+0.07}_{-0.06}$	$1.57^{+0.19}_{-0.10}$
Fornax (Fnx)	138	0.39	2.70	15.5	$0.12^{+0.07}_{-0.04}$	$1.14^{+0.09}_{-0.12}$
Leo II	205	0.19	0.52	0.58	$0.16^{+0.03}_{-0.07}$	$1.43^{+0.23}_{-0.15}$
Leo I	250	0.20	0.80	4.79	$0.06^{+0.14}_{-0.01}$	$1.45^{+0.27}_{-0.20}$
Segue 1 (Seg 1)	25	0.031	–	3.4×10^{-4}	$0.35^{+0.58}_{-0.24}$	$1.58^{+3.30}_{-1.11}$
Ursa Major II (Uma II)	32	0.127	–	4.0×10^{-3}	$0.31^{+0.18}_{-0.10}$	$1.09^{+0.89}_{-0.44}$
Willman 1 (W1)	38	0.025	–	1.0×10^{-3}	$0.23^{+0.18}_{-0.09}$	$0.77^{+0.89}_{-0.42}$
Coma (Com)	44	0.064	–	3.7×10^{-3}	$0.19^{+0.09}_{-0.05}$	$0.72^{+0.36}_{-0.28}$
Ursa Major I (Uma I)	106	0.308	–	1.4×10^{-2}	$0.34^{+0.15}_{-0.09}$	$1.10^{+0.70}_{-0.29}$
Hercules (Her)	138	0.321	–	3.6×10^{-2}	$0.19^{+0.10}_{-0.07}$	$0.72^{+0.51}_{-0.21}$
CV II	151	0.132	–	7.9×10^{-3}	$0.19^{+0.14}_{-0.07}$	$0.70^{+0.53}_{-0.25}$
Leo IV	158	0.152	–	8.7×10^{-3}	$0.12^{+0.14}_{-0.09}$	$0.39^{+0.50}_{-0.29}$
CV I	224	0.554	–	2.3×10^{-1}	$0.34^{+0.20}_{-0.08}$	$1.40^{+0.18}_{-0.19}$
Leo T	417	0.170	–	5.9×10^{-2}	$0.39^{+0.25}_{-0.13}$	$1.30^{+0.88}_{-0.42}$

Table 1: Table of properties of Milky Way satellites. Where shown, in parenthesis we list the abbreviation used in Figure 2. Dwarf satellites listed below the horizontal dividing line denote the newly-discovered population of SDSS dwarfs. For the classical satellites, we use King profiles, described by the King core radius, r_{king}. For the new satellites, we use Plummer profiles, described by r_{pl}. The limiting radius, r_{lim}, is defined as the limiting radius for the King profile. The error bars on the mass values reflect the points at which the likelihood function falls off to 60.6% of its peak value on either side. The luminosities for the newly-discovered satellites are taken from Ref. [47].
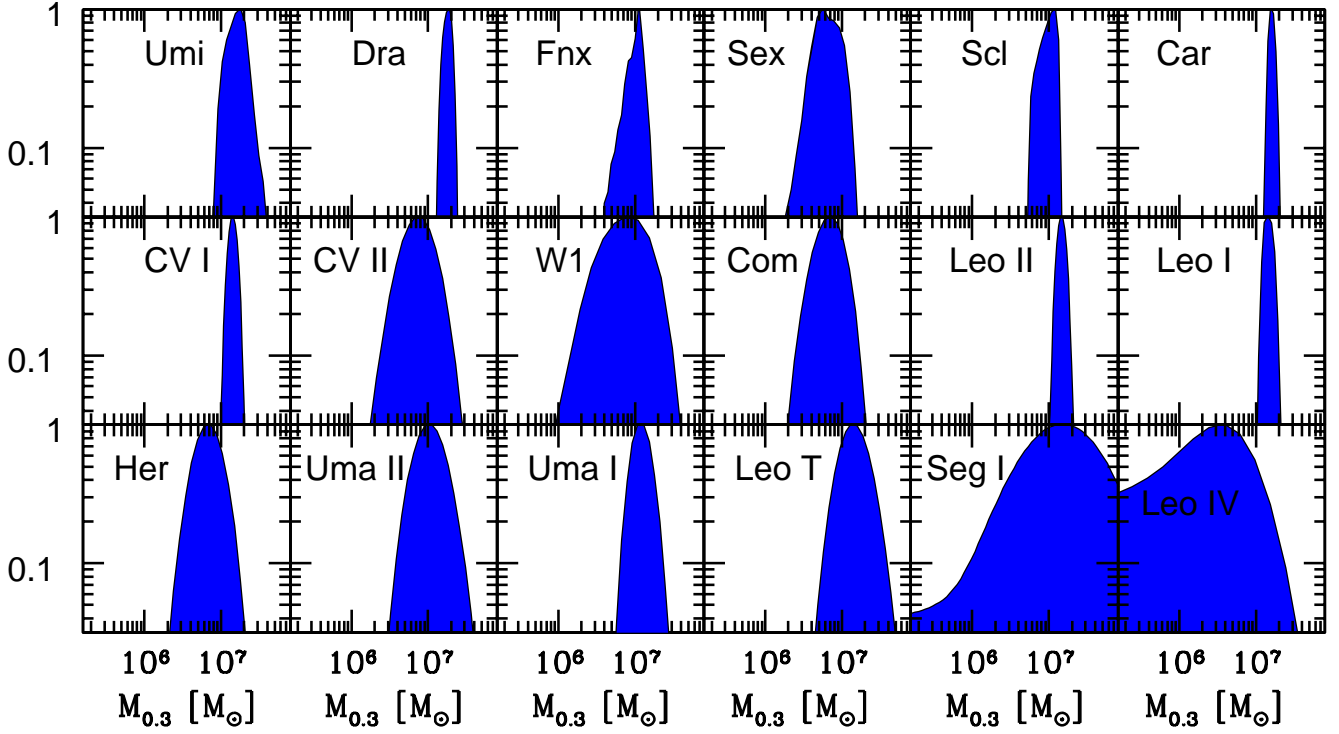
Figure 2: The likelihood function for the integrated mass within 0.3 kpc for 18 of the Milky Way satellites. We marginalize over all parameters as described in the text.
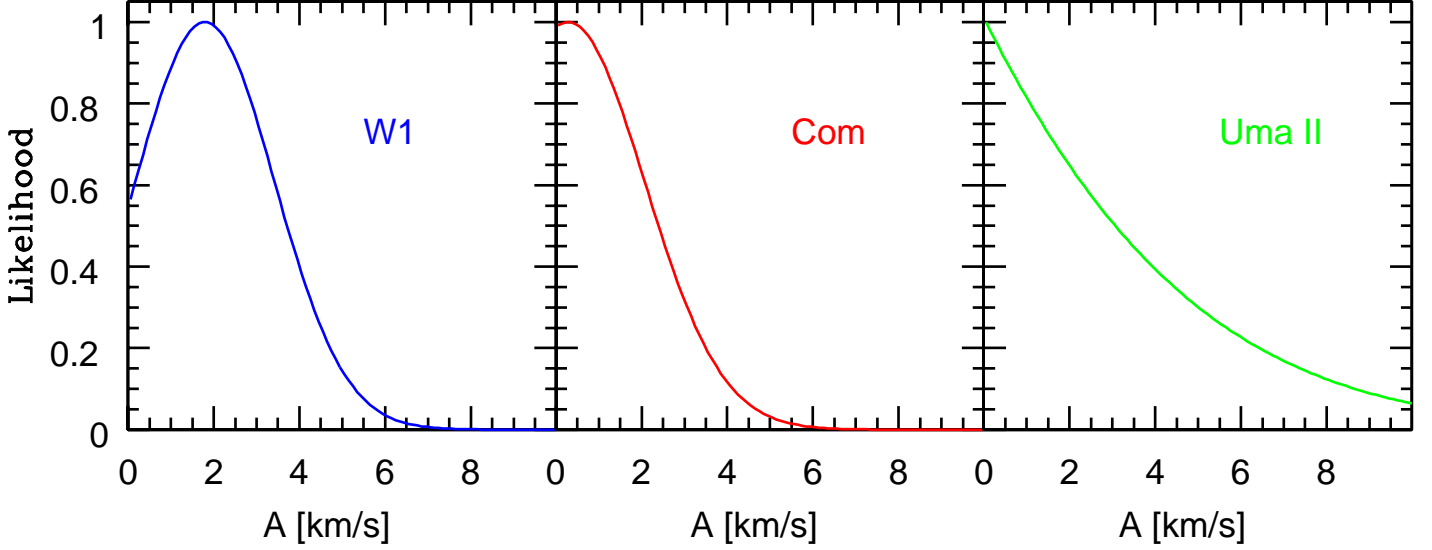
Figure 3: The likelihood function for the rotation amplitude for Willman 1, Coma Berenices, and Ursa Major II.

1. Willman, B. *et al.* A New Milky Way Companion: Unusual Globular Cluster or Extreme Dwarf Satellite? *Astron. J.* **129**, 2692–2700 (2005).

2. Belokurov, V. *et al.* Cats and Dogs, Hair and a Hero: A Quintet of New Milky Way Companions. *Astrophys. J.* **654**, 897–906 (2007).

3. Mateo, M. L. Dwarf Galaxies of the Local Group. *Ann. Rev. Astron. Astrophys.* **36**, 435–506 (1998).

4. Gilmore, G. *et al.* The Observed Properties of Dark Matter on Small Spatial Scales. *Astrophys. J.* **663**, 948–959 (2007).

5. Spergel, D. N. *et al.* Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. *Astrophys. J.* **170**, S377–S408 (2007).

6. Simon, J. D. & Geha, M. The Kinematics of the Ultra-faint Milky Way Satellites: Solving the Missing Satellite Problem. *Astrophys. J.* **670**, 313–331 (2007).

7. Walker, M. G. *et al.* Velocity Dispersion Profiles of Seven Dwarf Spheroidal Galaxies. *Astrophys. J.* **667**, L53–L56 (2007).

8. Peebles, P. J. E. Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. *Astrophys. J.* **263**, L1–L5 (1982).

9. White, S. D. M., Frenk, C. S. & Davis, M. Clustering in a neutrino-dominated universe. *Astrophys. J.* **274**, L1–L5 (1983).

10. Blumenthal, G. R., Faber, S. M., Primack, J. R. & Rees, M. J. Formation of galaxies and large-scale structure with cold dark matter. *Nature* **311**, 517–525 (1984).

11. Klypin, A., Kravtsov, A. V., Valenzuela, O. & Prada, F. Where Are the Missing Galactic Satellites? *Astrophys. J.* **522**, 82–92 (1999).

12. Moore, B. *et al.* Dark Matter Substructure within Galactic Halos. *Astrophys. J.* **524**, L19–L22 (1999).

13. Diemand, J., Moore, B. & Stadel, J. Earth-mass dark-matter haloes as the first structures in the early Universe. *Nature* **433**, 389–391 (2005).

14. Diemand, J., Kuhlen, M. & Madau, P. Dark matter substructure and gamma-ray annihilation in the Milky Way halo. *Astrophys. J.* **657**, 262 (2007).

15. Bode, P., Ostriker, J. P. & Turok, N. Halo Formation in Warm Dark Matter Models. *Astrophys. J.* **556**, 93–107 (2001).

16. Efstathiou, G. Suppressing the formation of dwarf galaxies via photoionization. *Mon. Not. R. Astron. Soc.* **256**, 43P–47P (1992).

17. Kauffmann, G., White, S. D. M. & Guiderdoni, B. The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. *Mon. Not. R. Astron. Soc.* **264**, 201–+ (1993).

18. Bullock, J. S., Kravtsov, A. V. & Weinberg, D. H. Reionization and the Abundance of Galactic Satellites. *Astrophys. J.* **539**, 517–521 (2000).

19. Kravtsov, A. V., Gnedin, O. Y. & Klypin, A. A. The Tumultuous Lives of Galactic Dwarfs and the Missing Satellites Problem. *Astrophys. J.* **609**, 482–497 (2004).

20. Mayer, L., Kazantzidis, S., Mastropietro, C. & Wadsley, J. Early gas stripping as the origin of the darkest galaxies in the Universe. *Nature* **445**, 738–740 (2007).

21. Muñoz, R. R. *et al.* Exploring Halo Substructure with Giant Stars: The Dynamics and Metallicity of the Dwarf Spheroidal in Boötes. *Astrophys. J.* **650**, L51–L54 (2006).

22. Martin, N. F., Ibata, R. A., Chapman, S. C., Irwin, M. & Lewis, G. F. A Keck/DEIMOS spectroscopic survey of faint Galactic satellites: searching for the least massive dwarf galaxies. *Mon. Not. R. Astron. Soc.* **380**, 281–300 (2007).

23. Strigari, L. E. *et al.* Redefining the Missing Satellites Problem. *Astrophys. J.* **669**, 676–683 (2007).

24. Mateo, M., Olszewski, E. W., Pryor, C., Welch, D. L. & Fischer, P. The Carina dwarf spheroidal galaxy - How dark is it? *Astron. J.* **105**, 510–526 (1993).

25. Piatek, S. & Pryor, C. The effect of galactic tides on the apparent mass-to-light ratios in dwarf spheroidal galaxies. *Astron. J.* **109**, 1071–1085 (1995).

26. Fellhauer, M. *et al.* Is Ursa Major II the Progenitor of the Orphan Stream? *Mon. Not. R. Astron. Soc.* **375**, 1171–1179 (2007).

27. Dunkley, J. *et al.* Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data (2008). `arXiv:0803.0586[astro-ph]`.

28. Bullock, J. S. *et al.* Profiles of dark haloes: evolution, scatter and environment. *Mon. Not. R. Astron. Soc.* **321**, 559–575 (2001).

29. Dekel, A. & Silk, J. The origin of dwarf galaxies, cold dark matter, and biased galaxy formation. *Astrophys. J.* **303**, 39–55 (1986).

30. Wyithe, J. S. B. & Loeb, A. Suppression of dwarf galaxy formation by cosmic reionization. *Nature* **441**, 322–324 (2006).

31. Walker, M. G. *et al.* Internal Kinematics of the Fornax Dwarf Spheroidal Galaxy. *Astron. J.* **131**, 2114–2139 (2006).

32. Koch, A. *et al.* Stellar Kinematics in the Remote Leo II Dwarf Spheroidal Galaxy-Another Brick in the Wall. *Astron. J.* **134**, 566–578 (2007).

33. Kuhlen, M., Diemand, J. & Madau, P. The Shapes, Orientation, and Alignment of Galactic Dark Matter Subhalos. *Astrophys. J.* **671**, 1135–1146 (2007).

34. Binney, J. & Tremaine, S. *Galactic dynamics* (Princeton, NJ, Princeton University Press, 1987, 747 p., 1987).

35. King, I. The structure of star clusters. i. an empirical density law. *Astron. J.* **67**, 471 (1962).

36. Irwin, M. & Hatzidimitriou, D. Structural parameters for the Galactic dwarf spheroidals. *Mon. Not. R. Astron. Soc.* **277**, 1354–1378 (1995).

37. McConnachie, A. W., Peñarrubia, J. & Navarro, J. F. Multiple dynamical components in Local Group dwarf spheroidals. *Mon. Not. R. Astron. Soc.* **380**, L75–L79 (2007).

38. Klimentowski, J. *et al.* Mass modelling of dwarf spheroidal galaxies: the effect of unbound stars from tidal tails and the milky way. *Mon. Not. R. Astron. Soc.* **378**, 353–368 (2007).

39. Illingworth, G. The masses of globular clusters. II - Velocity dispersions and mass-to-light ratios. *Astrophys. J.* **204**, 73–93 (1976).

40. Walsh, S. M., Jerjen, H. & Willman, B. A Pair of Boötes: A New Milky Way Satellite. *Astrophys. J.* **662**, L83–L86 (2007).

41. Ibata, R. A., Wyse, R. F. G., Gilmore, G., Irwin, M. J. & Suntzeff, N. B. The Kinematics, Orbit, and Survival of the Sagittarius Dwarf Spheroidal Galaxy. *Astron. J.* **113**, 634–655 (1997).

42. Majewski, S. R. *et al.* Exploring Halo Substructure with Giant Stars. II. Mapping the Extended Structure of the Carina Dwarf Spheroidal Galaxy. *Astron. J.* **119**, 760–776 (2000).

43. Merritt, D. The distribution of dark matter in the coma cluster. *Astrophys. J.* **313**, 121–135 (1987).

44. Drukier, G. A. *et al.* Global kinematics of the globular cluster M15. *Astron. J.* **115**, 708–+ (1998).

45. Feast, M. W., Thackeray, A. D. & Wesselink, A. J. Analysis of radial velocities of stars and nebulae in the Magellanic Clouds. *Mon. Not. R. Astron. Soc.* **122**, 433–+ (1961).

46. Olszewski, E. W., Pryor, C. & Armandroff, T. E. The Mass-to-Light Ratios of the Draco and Ursa Minor Dwarf Spheroidal Galaxies. II. The Binary Population and its Effects on the Measured Velocity Dispersions of Dwarf Spheroidals. *Astron. J.* **111**, 750–+ (1996).

47. Martin, N. F., de Jong, J. T. A. & Rix, H.-W. A comprehensive Maximum Likelihood analysis of the structural properties of faint Milky Way satellites (2008). `arXiv:0805.2945[astro-ph]`.